VEGETALJUCE'S GIGANTIC GUIDE TO GUMI ENGLISH

Intro to Gumi English Pronunciation……...Page 1

Phoneme list……..…….Page 2

Stops, chops, and breaks……...Page 5

Breaks between words……..P. xxx

Breaks inside words……..……..…….P. xxx

Choppy L's……...……...P. xxx

Pronunciation issues……...P. xxx

Fuck what the dictionary says…….P. xxx

Abnormal vowel pronunciation…….P. xxx

Soft Consonants……..…….P. xxx

Soft Vowels…….……..…….P. xxx

Compound phoneme mispronunciation…….P. Xxx

Taking your pronunciation to the next level…….……P. xxx

Log of specific fixes……..……..……...P. xxx

* Blue represents my progress so far

INTRO TO GUMI ENGLISH PRONUNCIATION

Gumi is widely considered to have the best english voice bank out of all the non-native english vocaloids, and for good reason. The samples that her voice provider, Megumi Nakajima, provided are all pronounced with very minimal accent. The sounds she has are all great, the problem stems, I believe, from the fact that the programming that puts them all together into a cohesive voicebank is somewhat shotty.

The most common issues you'll encounter when using her voice bank are **broken transitions** (eg. choppiness inside or between words), and **inconsistent vowel pronunciation**. These issues in turn are affected by things such as the **pitch** they are on or the **phonemes around them**. The good news is that nearly all these problems can be remedied.

Over the course of several years, I've figured out a number of fixes and workarounds for her voicebank. My hope is that this guide can save those who read it countless hours of trial and error testing by laying out all the tips and tricks I've gathered right here.

Before we get started, here are the basics you need to know about vocaloid pronunciation.

- When you type in a word, Vocaloid will spit out a default phonetic spelling, seen in [brackets].
- This default phonetic spelling is not always the best for your vocaloid, and can be changed by editing the phonemes.
- You can switch from word editing to phoneme editing by pressing CRTL+R.
- Phonemes are represented by letters and symbols, and are different depending on the language of the voicebank you're using. It is helpful to know them all, and you'll wind up learning them over time. A complete list of English phonemes can be found on the next pages.

If you can't find the answer to a problem you're having in this guide, feel free to contact me with your question at wrencovers@gmail.com and I'll do my best to help you out.

ENGLISH VOCALOID PHONEMES (& HOW THEY SOUND WITH GUMI ENGLISH)

VOWELS

PHONEME	EXAMPLES AND NOTES
V	What, Abort *Slightly "sharper" but very similar to [Q]*
Q	Pond, Body, Was. *Slightly less "harsh" than [V] but very similar.*
@	*Short, throaty "ah" sound. Subtle. Usually sounds weird on its own. Useful for words like "little" as the vowel before the ending L.*
{	Cat, Apple, Plan *Often the transition between a consonant and this phoneme is broken. Can also sound a little soft at times.*
e	Met, When, Bed
I	Dinner, Pick, Tin *When held for too long can start sounding like [i:]*
i:	Me, Sweet, Queen
U	Sun, But, Up
u:	Spoon, Crew, Too
O:	All, Nod *Can also be used to replace [V] or [Q] if they aren't working right.*
@U	Toe, Crow, Know
aU	Out, Cow, Pound
aI	Tie, Cry, Fight
eI	May, Ate, Aim
OI	Choice, Toy, Moist

CONSONANTS

PHONEME	EXAMPLES AND NOTES
b OR bh *	Bad, Abs, Bus
d OR dh *	Doll, Mud, Drag
f	Face, Life, Fool
g OR gh *	Ghost, Bag, Gap
h	Hide, Who, Ham
dZ	Jump, Badge, Jam
Dz	Pleasure, measure, azure
k OR kh *	Car, Take, Cross
l OR l0	Lie, Feel, Lack *Depending on the phonemes around it, the default phonetic spelling of a word will use one or both of these. See pages 8 & 9 for more details.*
m	Man, Home, Mouse
n	Neck, Can, Knit
N	Long, Sing, Bang
p OR ph *	Peach, Cap, Pale
r	Rise, Room, Run
s	Sun, Pass, Swing
S	Shape, Ash, Shoot
tS	Choice, Patch, Chair
T OR th *	Trim, Mint, Tame
v	Viral, Save, Victim
w	When, Wake, War
j	Yes, You, Your

z	Zoom, Zebra, Zip

*** Usually the one with the H is for the beginning of a word and the plain one is for the end of a word, but if one is sounding bad, try the other.*

“COMPOUND” PHONEMES

PHONEME	EXAMPLES AND NOTES
@r	Curse, World, Sure
e@	Air, Bear, Hair
Q@	Are, Star, Car
I@	Beer, Here, Tier

EXTRA PHONEMES

PHONEME	EXAMPLES AND NOTES
-	Called a portamento. Carries a sound through all of the subsequent notes that have this on them.
Sil	Stops sound short
_0	Tack this on the end of a phoneme for a whispered version

STOPS, CHOPS, AND BREAKS

A frequent complaint about Gumi English is that she sounds “choppy”. The choppiness is caused by her many broken transitions. A **transition** is the ability of one phoneme to smoothly blend into the next. When there is a stop between two phonemes instead of a smooth blend, this is called a “broken” transition.

In order to fix this, you have to bridge the gap in the transition. The solution for this differs depending on the place a broken transition falls and the phonemes involved.

BROKEN TRANSITIONS BETWEEN WORDS

Breaks between words are usually the easiest to fix. This most commonly occurs when a word in the middle of a phrase starts with [aI] or [V]. Take for example, the phrase “Blue eyes”:

[bh l u:]⇎[aI z]

In this case, there is a broken transition between the [u:] and the [aI] phoneme. This is due to the fact that vowel phonemes generally have trouble transitioning into [aI]. Because we can’t replace either of those phonemes with something different, we have to find a way to “bridge the gap”. That is, we have to find a phoneme to put between them that has an intact transition to both [u:] and [aI]. Of course, it has to be one that doesn’t sound out of place there. As luck would have it, there’s a perfectly natural choice. If you say “blue eyes” out loud, you’ll notice there’s a sort of “w” sound between the two words. In order to fix this transition, we need only to insert that missing sound!

[bh l u:]⇔[w aI z]

(You can put the [w] at the end of blue or the beginning of eyes, but I tend to place non-vowel bridges like [w] or [j] at the beginning of the second word so I can control them through the VEL parameter if need be.)

This same fix can be used if the phrase was “Grey eyes”, only instead of the [w] , you’d use a [j] instead. When there’s a break between words where the offending phonemes are two vowels, you will always use either [w] or [j].

Another time you'll run into this issue is when the previous word ends with a compound phoneme (see page 4). Take the phrase "Their eyes" for example:

[D e@]⇎[aI z]

In this case there's a broken transition between [e@] and [aI]. If you wanted to fix this one in the same way we did the last one, the solution would be to keep "their" as it is and put an [r] at the beginning of "eyes". However, while this mends the break, it will also put a bit too much emphasis on the "R". That's because technically, we're using two "R" sounds (both [e@] and then [r]). So we need to get rid of one. We know [e@] is causing us problems, so it's better to get rid of that one. Of course, because [e@] is a compound phoneme, when we get rid of it, we're also getting rid of the vowel, so we have to replace it. You might be tempted to just use [e], but if you say the word "their" out loud and stop just before you say the "R", you'll notice the vowel should actually be [eI]. Completed, this is what the fix should look like:

[D eI]⇔[r aI z]

SUMMARY

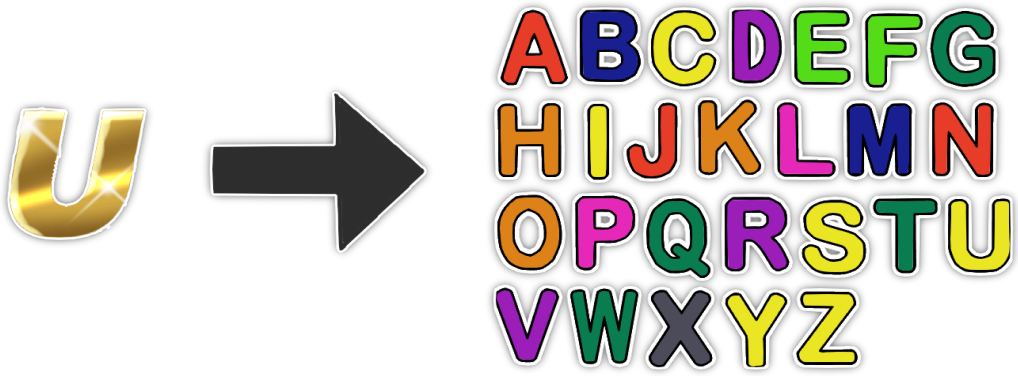
When there's a break between words caused by two vowels, use either [w] or [j]. When the break is caused by a compound phoneme, replace the compound phoneme with the appropriate vowel and add an [r] to the second word

BROKEN TRANSITION INSIDE A WORD

A break inside a word can be a little more difficult to fix. This issue most commonly arises with the [s], [S] or [tS] phonemes. When a vowel follows one of these, there is often a break. For example, take the word "choice":

[tS ⇎ OI s]

Here, the [tS] doesn't transition into [OI] at all. The result is a very ugly, glitchy sounding chop. But fear not, for we can bridge this gap with what I like to call "the magic [U] phoneme".

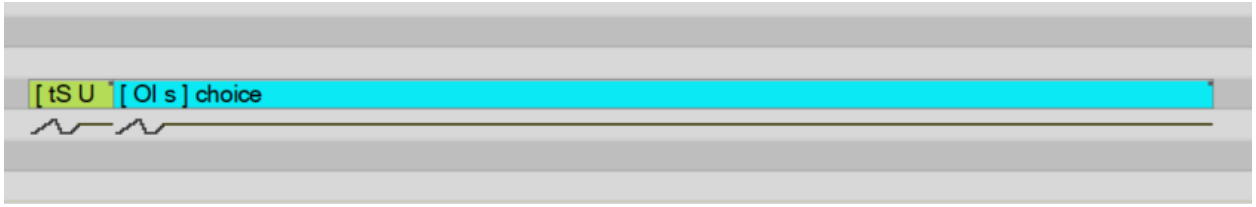


The [U] phoneme is so handy because it has an intact transition to almost every other phoneme, which although commonplace in most other voicebanks, is something special indeed when it comes to Gumi English. Not only this, but it has a mild sound that won't throw off the pronunciation of the word too much when inserted.

This makes [U] essentially a perfect band-aid to almost every internal broken transition. So let's apply it to our example:

[tS⇔U⇔OI s]

Perfect… Almost. You can technically leave it like this if you play it back and it sounds okay, but sometimes you may notice that the [U] is getting pronounced for too long and is skewing the pronunciation a little. Time to send it into **quarantine**:



To do this, split the word in two right after [U], and make the first note (the one containing the tS and U) much shorter than the other half. I typically do 1/64th, aka the smallest length possible without turning off length mode. This constrains the amount of time [U] is pronounced to almost nothing, which keeps the original vowel pronunciation intact.

SUMMARY

When there's a break inside a word, you can almost always use [U] to mend the broken transition by inserting it between the two offending phonemes.

CHOPPY L'S

Many times you'll find that Gumi's L's sound choppy. If left unfixed, this can seriously affect her understandability and the phrase's overall smoothness. When you encounter this issue, you can always start by swapping out one L phoneme for the other. For instance, if the default phonetic spelling of a word uses [l0], you can try replacing it with [l] and vise versa. This often does the trick if the L falls at the beginning or end of a phrase, but if the L is within a phrase with phonemes on either side, this quick fix won't always work.

Again, it comes down to broken transitions. Both L phonemes weren't created equal in regards to what transitions they have. [l] might transition into one phoneme just fine, while [l0] will fail to transition into that very same phoneme. This is why, when an L falls between phonemes, there will be cases where neither of the L phonemes work on their own, since the one that has an intact transition with the phoneme before it might not have an intact transition with the phoneme after it. In this case, even if you swap it for the other L phoneme, the new L phoneme will have the same problem only reversed. The best way to fix this problem is to use **both** L phonemes.

Typically, I just put them both next to each other like this [l0 l] or like this [l l0]. Depending on

the phonemes around them, one of the two orders will work and one of them won't, so you'll have to take it on a case by case basis.

SUMMARY

There are two L phonemes, [l] & [l0]. When one sounds choppy, try swapping it for the other. If that doesn't work, use them both together (you may need to switch the order).

PRONUNCIATION ISSUES

This section focuses on general pronunciation issues that are not related to **stops, chops, or breaks**. This section, unlike the previous, will not feature concrete fixes or rules. This is due to the fact that there are nearly infinite amounts of phoneme combinations that could result in mispronunciations (so many that I run into new ones every day despite using Gumi English for years!) Instead, this section is meant to act more like a tool kit of concepts and ways of thinking about the issues you encounter.

FUCK WHAT THE DICTIONARY SAYS

As I'm sure you've noticed, when you're not in phoneme entering mode and you type in a word the dictionary recognizes, vocaloid will spit out a default phonetic spelling of that word. A lot of the time this spelling is just fine, but sometimes vocaloid will give you a spelling that doesn't sound right with the voicebank you're using.

With Gumi English, **the issue will most frequently be with the vowel the dictionary has chosen.** For example, take the word "fun". The default spelling of this word is as follows:

[f V n]

With this spelling, the [V] makes the word sound more like *"fan"*. In order to fix this, we can swap the [V] with a different vowel phoneme from within its same sound group. Below is a chart showing these **groups of similar vowel phonemes.**

V Q O: U @*	I i	u: U	e {

Now it's just down to trial and error as you swap in the other phonemes to find out which is the right one. In this case, the best phoneme is actually [Q], so the best phonetic spelling is:

[f Q n]

ABNORMAL VOWEL PRONUNCIATION

Sometimes words use the technically correct vowel phoneme, yet it still comes out sounding wrong. Take, for example, the phonetic spelling of the word “caught”:

[kh Q t]

If you’re familiar with how Gumi’s vowel phonemes are supposed to sound, then you’ll know that this spelling, by all logic, should sound correct. Yet, when you play it back, you’ll find that the word sounds like “coat” instead of “caught.” Why is [Q] taking on a hard “O” sound? The answer lies in transitions once again.

Wonky transitions skew the pronunciation of phonemes, making them sound different than they’re supposed to, which is what’s going on here. Since it’s in the middle of a word, this [Q] has two transitions— one of which has to be the culprit. To narrow down which transition is at fault, start experimenting. First try changing the phoneme on the left to a different, random consonant, and listen for a change in the pronunciation of the vowel. If there’s no change, try swapping the consonant on the right and listening again. Changing one of the two phonemes on either side of the problematic vowel should yield the “proper” vowel pronunciation. The one whose replacement fixes the vowel sound is the one that has a wonky transition with the vowel.

Now that you’ve located the wonky transition, you need to figure out a way to change or interrupt it. For instance, in the example word [kh Q t], let’s say the wonky transition is between the [kh] and the [Q]. In this case, here are some potential fixes you might come up with:

1. Change the first phoneme to [k]
2. Insert a magic “U” between the first phoneme and the [Q]
3. Change the [Q] to a [V] or an [O:]

If the wonky transition is between the [Q] and the [t], here are some fixes you might try:

1. Change the [t] to a [th]
2. Insert a magic “U” between the [Q] and the [t]
3. Insert a [dh] or a [d] before the [t]

SOFT CONSONANTS

Sometimes, usually at the beginning or end of a word, a consonant will be much softer than it should be. At times, it might even be nearly inaudible. Take, for example, the word “brain”:

[bh r aI n]

Here, the “B” sound is so soft that the word almost sounds like “rain”. As always, this is a transition issue, this time between [bh] & [r]. However, in this case, we are somewhat limited in fix options due to the fact that there is an [r] involved. [r] can be an annoying phoneme to deal with because it has no “alternative” phoneme that it can be swapped out for since it’s the only one that can make a stand-alone “R” sound. So, of course, we can’t change it-- even though it’s part of what’s causing the wonky transition. This means the only option is to change the [bh].

1. [b]
2. [p]
3. [ph]

LOG OF SPECIFIC FIXES

This is where I'll document fixes I've come up with for specific words. They are listed in alphabetical order. I'll update the list with more when I can.

**Keep in mind that sometimes fixes are limited to a certain note-range; depending on the pitch the word is on, you may find that it actually doesn't need to be fixed, or you may find that the listed fix doesn't work in that range. That said, I suspect that as long as you are working in her mid-range, you probably won't run into this phenomenon too often.*

BLUE = Long note
YELLOW = Mid-length note
RED = Short note (1/64)

WORD	ISSUE	FIX
Blank	Broken transition, soft consonant, wonky vowel	[bh p l0 e { N k]
Breathe	Wonky vowel	[bh r i: l] [i: D]
Down	Weak vowel	[d {] [w l] [n]
Fire	Weak R	[f Q] [j @r]
Gold*	Weak vowel, mispronounced L	[gh @U u: l0 l th d]
World*	Wonky vowel, mispronounced L	[w U @r l l0 dh]